

## DOCUMENT RESUME

ED 199 229

SP 017 646

AUTHOR Hales, Gary  
TITLE Measuring Moral Judgement: The Implications of Situation Specificity.  
PUB DATE 81  
NOTE 15p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Dallas, TX, January 29-31, 1981).

EDRS PRICE MF01/PC01 Plus Postage.  
DESCRIPTORS \*Affective Measures; \*Attitude Measures: Decision Making: Developmental Stages; \*Ethics: Human Relations; \*Moral Values; Opinions: Postsecondary Education: Secondary Education; Social Attitudes; \*Social Problems; \*Student Attitudes

## ABSTRACT

Research has suggested that moral judgment is a construct which can be measured to some degree by an objectively scored instrument. This study investigated whether moral judgment is stable enough to be accurately measured. A test of moral judgment was developed, and special care was taken to ensure that the researcher was aware of the moral issues to which the participants were responding. Students, ranging from junior high school level to post-college, participated. Moral judgments made with reference to the pairs of moral issues presented in two dissimilar contexts were compared. Findings suggested that moral judgment may be situation specific. To the extent that moral judgment is situation specific, accurate measurement would be difficult, because there is an infinite number of scenarios in which any two moral issues could be portrayed. (Author/JD)

\*\*\*\*\*  
\* Reproductions supplied by EDRS are the best that can be made \*  
\* from the original document. \*  
\*\*\*\*\*

GARY HALEG  
U.T. AUSTIN  
1981

Paper presented at the annual meeting of the  
Southwest Educational Research Association  
(Dallas, 1981).

MEASURING MORAL JUDGEMENT: THE IMPLICATIONS OF SITUATION SPECIFICITY

ABSTRACT

Previous research in moral judgment has suggested that it is a construct which can be measured to some degree by an objectively scored instrument; for example, Rest's Defining Issues Test (DIT). The research question investigated in this study was whether moral judgment is stable enough to be accurately measured. A test of moral judgment, similar in format to the D.I.T., was developed and special care was taken to ensure that the researcher was aware of the moral issues to which the student participants were responding. Subsequently, 395 students, ranging in educational level from junior high school level to post-college, responded to the instrument. The moral judgments made with reference to the pairs of moral issues presented in two dissimilar contexts were compared: there were 12 moral dilemmas in all. Results were analyzed using a repeated measures analysis of variance.

The findings suggest that moral judgment may in fact be situation specific. To the extent that moral judgment is situation specific accurate measurement would be difficult since there are an infinite number of scenarios in which any two moral issues could be portrayed.

U.S. DEPARTMENT OF EDUCATION  
NATIONAL INSTITUTE OF EDUCATION  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

- ☒ This document has been reproduced as  
received from the person or organization  
originating it.  
☐ Minor changes have been made to improve  
reproduction quality.

• Points of view or opinions stated in this docu-  
ment do not necessarily represent official NIE  
position or policy.

"PERMISSION TO REPRODUCE THIS  
MATERIAL HAS BEEN GRANTED BY

Gary D. Hales

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC)."

ED199229

SP 017646

## REVIEW OF THE LITERATURE

The initial impetus for work in studying moral judgment comes from psychological perspectives suggested by developmental theorists, e.g. Piaget. The underlying assumption that moral judgment develops in stages is certainly closely related to Piagetian cognitive development stages, i.e., moral judgment phases build upon one another but are qualitatively different from each other.

Piaget postulated two primary stages of moral development: 1) heteronomous morality (H.T.); and, 2) autonomous morality (A.T.) (Lickona, 1976; Ausubel & Sullivan, 1970). Lickona differentiated the two stages on nine dimensions; for example, absolutism of moral perspective represented H.T. while awareness of differing viewpoints represented A.T. Piaget measured sophistication in moral judgment by telling a child a story about a moral dilemma and then asking the child about the characters in the story and the action these characters might take in response to a particular situation.

Piagetian investigation of moral judgment was concerned primarily with the ability of the child to make moral decisions based on information which is related to either intentions or to consequences. Research reviewed by Lickona (1976) showed mounting evidence of within area (of moral judgment) variability of a child's judgment.

A major force not only in measurement of moral judgment but in the scientific study of construct comes from Lawrence Kohlberg's work. Kohlberg postulated six stages of moral development that a person might progress through. He felt that a moral judgment stage was a structured

whole and represented an underlying pattern of thought organization (1963,1964,1969). An individual was tested to determine what stage he/she was in by presenting a moral dilemma to the subject and then asking the person what the main character would do and why (Kohlberg, 1958, 1973). Findings from these testings which are most relevant to situation specificity are:

1. An invariant irreversible sequence of moral development has not been verified.
2. Greater stage unity was found when moral dilemmas developed by Kohlberg rather than by Piaget were used (both from Simpson 1974).

One of the criticisms leveled at Kohlberg's work has been that his theoretical approach suggests that an individual judges from one particular stage, at a particular point in time.

Rest felt that an individual respondent might use different stage typical moral concepts at different times. Rather than stage typing an individual his testing procedures produced a score indicating the percentage of use of principled moral concepts. The Defining Issues Test (Rest, 1975) (DIT) was the objectively scored instrument he used to measure moral judgment. Rest (1973) found that subjects preferred moral stage statements at the highest level they could comprehend. Results suggested that individuals did not think at one stage 100 percent of the time; there were gradual shifts in stage thinking which resulted in persons considering a problem from more than one stage perspective as their thinking developed in logical consistency and explicitness. Rest (personal communication 1976) suggested the

possibility of situation specificity of moral judgments and encouraged the development of moral dilemma stories to more adequately sample the domain of moral problems.

Work by Rest suggested that perhaps the individual when making moral judgments does so based not only on the moral issues involved but also on the basis of the type of situation in which the moral dilemma takes place. If this reflects the actual state of affairs, then one must question the advisability of attempting to measure moral judgment. Since there are an infinite set of scenarios in which a moral dilemma could take place it would not be possible to obtain an accurate measure of an individual's moral judgment sophistication without testing that sophistication in all possible situations—clearly unfeasible. The primary purpose of the study was to investigate the question of situation specificity of moral judgment by use of an objective test instrument.

#### METHOD

Sample. Four hundred seventy students, ranging in educational level from junior high school (grades 7 & 8) to law students with baccalaureate degrees, were asked to respond to the DecisionsI test. After checking for internal consistency of the test instrument by using procedures suggested by Rest (1974) and additional procedures developed by Hales (1979) a total of 395 students were included in the data analysis. Participation was voluntary and all subjects signed a consent form; subjects under the age of 18 had consent forms signed by their parent or guardian.

Instrumentation. The test instrument contained twelve moral

dilemma stories juxtaposing four moral issues in all possible pairs. The moral issues were: human life, individual rights, majority rights, and material good. Material good included such things as great art, unspoiled nature, etc., which were designed to attribute value to entities in the absence of human involvement. The stories were field tested prior to use in the study to ensure that the majority of respondents felt the dilemma stories were portraying the moral issues the experimenter had written them to portray. This was essential since the responses would be evaluated on the basis that the respondents were judging about the same moral issues simply described in different situations. Two slightly different test forms were used, the major difference being that one set of students were asked to choose moral statements they thought were important and agreed with, while another smaller group chose statements that they thought were important. This modification was suggested by pilot testing.

In addition, the test was written at a seventh grade reading level so that all respondents would be able to understand the material involved. After examining the D.I.T., it was felt that part of the test was not moral sophistication but reading ability of the student which could mask sophistication of the respondent.

The instrument consisted of a 21-page questionnaire which presented the respondent with twelve moral dilemmas each followed by ten statements. Moral judgment stages 2 through 6 were each represented by two statements per story. Scoring was conducted as described in Rest (1974) and generated a P score (principled reasoning score) ranging from a possible low of 0 to a possible high of 120. The higher the P score the more sophisticated is the moral reasoning of the respondent as measured by the test.

Test/retest reliability of the instrument was .66 ( $n = 44$ ).

Testing. Testing was conducted in junior high schools and high schools of three small mid-central Texas cities and at The University of Texas at Austin. Testing of pre-college students was done in classroom setting while most college and law students completed the instrument at home. Instruments completed at the student's home were returned by means of a stamped, pre-addressed envelope. Since at-home conditions could not be controlled as closely as could testing in the classroom, instructions to the respondents for home use requested that they complete the test in one sitting, in quiet surroundings, while giving full attention to the task at hand.

Debriefing of pre-college students took place in the classroom, while most college and law students were debriefed at meetings held after the testing for the class group as a whole.

Hypothesis. The major hypothesis was:

There will be no significant difference in mean P scores for pairs of dilemmas written to juxtapose the same pair of moral issues.

For example, the dilemma storyscores for The Device were compared with dilemma story scores for the Trade, since these two stories both juxtaposed the moral issues of human life versus material good.

Analyses. A repeated measures analysis of variance was performed (using the ANOVA 8 Statistical Program, Marston, 1978). The results from these analyses are found in Table 1. The hypothesis of no statistically significant differences in average dilemma story scores between pairs of stories written to address the same two moral issues

Table 1

COMPARISON OF AVERAGE DILEMMA STORY SCORES BETWEEN PAIRS OF DILEMMA  
STORIES WHICH ADDRESS THE SAME TWO MORAL ISSUES

| Dilemma Story<br>Pair                          | Moral Issues                             | Form A     |            | Form B     |            |
|------------------------------------------------|------------------------------------------|------------|------------|------------|------------|
|                                                |                                          | F<br>Ratio | F<br>Prob. | F<br>Ratio | F<br>Prob. |
| Device/Trade                                   | Human life vs.<br>material good          | 15.786     | .000       | 2.296      | .134       |
| Chen/Teaching                                  | Individual rights vs.<br>majority rights | 3.120      | .078       | 2.545      | .115       |
| Attack/Farmer                                  | Human life vs.<br>individual rights      | 416.712    | .000       | 44.051     | .000       |
| For the Good of the Party/<br>Last Masterpiece | Material good vs.<br>individual rights   | 15.880     | .000       | 32.248     | .000       |
| Longhorn Lake/Energy<br>Crisis                 | Material good vs.<br>majority rights     | 1.617      | .204       | 3.998      | .045       |
| Voting about Lives/<br>Mountain Glen           | Human life vs.<br>majority rights        | 13.076     | .000       | .495       | .484       |

but in different contexts not supported in seven comparisons and supported in the remaining five comparisons. It may be noted that in three cases significant differences in average dilemma story scores were not found for both Form A and Form B simultaneously—Device/Trade, LonghornLake/EnergyCrisis, and Voting AboutLives/Mountain Glen. This is of interest because it suggests that differences in dilemma story scores could be influenced by which test form, either A or B, was used. The change in instruction between Forms A and B apparently did produce a noticeable effect in this case.

In just over half the comparisons the significant differences in average dilemma story scores was found comparing dilemma stories written to represent the same two moral issues, and this led to the conclusion that similarity in the moral judgment used to consider dilemmas with the same moral issues might not be as strong as was formerly thought.

#### CONCLUSIONS AND RECOMMENDATIONS

The hypothesis being tested was that when comparing moral dilemma stories which juxtaposed the same two moral issues in different context there would be no significant difference in moral judgment as indicated by a P score for each dilemma story. This hypothesis was supported in five of twelve comparisons made and not supported in the other seven comparisons. One explanation that could be offered for this finding is that dilemma stories with the same moral issue pairs were not in fact perceived as addressing the same moral issues. The experimenter is of the opinion that this does not take into account the care associated

with the development of the dilemma stories. An alternative explanation for the findings is that students are perceiving the same pair of moral issues in two dilemma stories and may also be reacting to other elements of the content. The student may, for example, find the story does not answer all the questions about the dilemma that he/she would like to know. The student might then extrapolate from given information and "fill out the story." It is readily apparent that such behavior would be very individualistic and could result in the respondent inserting into the story elements the experimenter had not intended. One method to address this would be to instruct the respondent to examine the story as if he/she were on a jury and had to make a decision based only on known facts, not on any opinions that he/she might have regarding undiscussed elements of the dilemma. It is not possible, of course, to provide a dilemma story which is so complete as to answer all questions of every respondent for then there would be no dilemma involved. The experimenter believes, however, that the above instructions could serve to standardize the stimulus material for the respondents on this dimension.

An additional problem is that while the dilemma story pairs may be correctly identified as addressing the same moral issue the content of the story may modify the impact of the moral issue. For example, the student might respond differently to the ATTACK than to the FARMER, because in the ATTACK story a character was allowed to die by a conscious decision not to act, i.e., not to help the men out of the car, while in the FARMER the death was not the result of a premeditated decision. In the construction of new dilemma stories care should be exercised so that the context is similar enough to lessen the effect of the problem discussed above.

The experimenter believes that the findings suggest that the story setting, or context of a dilemma story exerts influence on the moral judgments made about the content (moral issues involved). The results of the data analysis lend support to the hypotheses that moral judgment is at least somewhat situation specific; that is, will vary according to the particular situation regarding which the situation is made. If, however, a person's moral judgment is found to vary significantly between two situations involving the same moral issues the question would then arise, how measurable is moral judgment? One answer suggested by this study would be that the score generated by moral judgment test may be applicable only to certain moral issues and then only in specific situations found in that instrument. The measurement of moral judgment would be very difficult indeed if such judgment varied according to the situations that a moral dilemma occurred in. If moral judgment is found to vary according to the moral issues involved one could conceivably consult experts in philosophy and develop a rather comprehensive list of issues. There would be no way, however, to sample adequately all the possible situations in which such a list of issues stimulating moral dilemmas could arise, since there are an infinite number of possible situations for any one pair of moral issues. If moral judgment is found to be situation specific then valid measurement of moral judgment may not be an attainable goal. It is not the intention of the experimenter to state that moral judgment is so variable as to be unmeasurable. Given the twenty or more years of research in this area, given the results from this study did indicate no significant differences in five of the twelve comparisons, and given the suggested alternative explanations for the findings, such a conclusion would be premature. The experimenter feels, however, that

these results should lead other investigators to question what has appeared to be an a'priori assumption: judgment about the same moral issues in different situations should be stable.

In most cases the moral issues in the dilemma stories developed for use in other moral judgment tests are unstated at least in the references available to this experimenter. It is hoped that these findings will lead others to examine carefully the respondents' identification of what moral issues they—the respondents—feel the dilemma story is presenting. If further testing of moral judgment produces results similar to those found in this study the ability to measure moral judgment is even somewhat stable concept would need to be closely considered.

It is beyond the scope of this study to answer the question of "How and why does the situation in which moral dilemma takes place influence moral judgment?" The experimenter feels, however, that further research in the area of situation specificity of moral judgment is indicated and offers the following as a guide to such work. One method to investigate the effect of situation story setting is to use the same situation for a set of dilemma stories but write stories to present different moral issues. This would involve examining the same situation from different perspectives, that is, how does the situation appear when considered from the point of view of individual rights, majority rights, law and order, religion, etc. Dilemma story scores for each story could be compared in order to investigate the effect of story setting on moral judgment.

This study ought to be the beginning of a more careful investigation of moral judgments involving a variety of situations and moral issues. The experimenter would also suggest that once

instrumentation for the measurement of moral judgment has sufficient precision a greater effort be made to relate moral judgment to moral behavior.

# BIBLIOGRAPHY

- Ausubel, D. P. & Sullivan, E. V. Theory and problems of child development. New York: Grune & Stratton, 1972.
- Hales, G. D. Situation specificity of moral judgments manifested by a student population. Unpublished doctoral dissertation, The University of Texas at Austin, 1979.
- Kohlberg, L. The development of modes of moral thinking and choice in the in the years ten to sixteen. Unpublished doctoral dissertation, University of Chicago, 1958.
- Kohlberg, L. The development of children's orientation toward a moral: I. sequence in the development of moral thought. Vita Humana, 1963, 6, ii-33.
- Kohlberg, L. The development of moral character and ideology. In L. W. Hoffman (Ed.), Review of Child Development Research. New York: Russell Sage Foundation, 1964.
- Kohlberg, L. Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), Handbook of Socialization Theory and Research. New York: Rand McNally, 1969.
- Lickona, T. (Ed.). Moral development and behavior: Theory, research and social issues. New York: Holt, Rinehart, and Winston, 1976.
- Marston, P. T. ANOVA 8 manual: Instructions for using analysis of variance BMD08VA and orthogonal weighting program ORTH 2. Unpublished manual, Research and Development Center for Teacher Education, The University of Texas at Austin, 1978.
- Rest, J. The hierarchical nature of moral judgment: A study of patterns of comprehension and preference of moral stages. Journal of Personality, 1973, 41, 86-109.
- Rest, J. The cognitive-developmental approach to morality: The state of the art. Counseling and Values, 1974, 18, 64-78.
- Rest, J. New options in assessing moral judgment and criteria for evaluating validity. Paper presented at the Biennial meeting of the Society for Research in Child Development, Denver, April, 1975.
- Rest, J. Personal communication, December 17, 1976.